New convolutional code constructions and a class of asymptotically good time-varying codes

Justesen, Jørn

Published in:
I E E E Transactions on Information Theory

Publication date:
1973

Document Version
Publisher's PDF, also known as Version of record

New Convolutional Code Constructions and a Class of Asymptotically Good Time-Varying Codes

JØRN JUSTESEN

Abstract—**We show that the generator polynomials of certain cyclic codes define noncatastrophic fixed convolutional codes whose free distances are lowerbounded by the minimum distances of the cyclic codes. This result is used to construct convolutional codes with free distance equal to the constraint length and to derive convolutional codes with good free distances from the BCH codes. Finally, a class of time-varying codes is constructed for which the free distance increases linearly with the constraint length.**

I. Introduction

THIS paper contains a number of refinements and extensions of the results on the relationship between cyclic codes and convolutional codes given earlier by Massey *et al.* [1].

In Section II a lower bound is established on the free distance of certain convolutional codes derived from cyclic codes. The bound is given in terms of the minimum distance of the cyclic code and the roots of the generator polynomial.

In Section III convolutional codes analogous to the Reed–Solomon block codes are constructed, and in Section IV it is shown that the restriction of these codes to subfields produces good convolutional codes closely related to the BCH block codes.

Finally a technique quite similar to that used earlier by Justesen [2] to obtain asymptotically good block codes from Reed–Solomon codes is used to construct asymptotically good "periodic time-varying" convolutional codes from the convolutional codes of Section III.

II. A Lower Bound on the Free Distance of Certain Convolutional Codes

We shall make extensive use of the following theorem, which was proved by Massey *et al.* [1].

Theorem 1: For any polynomial $P(x)$ over $GF(p^r)$, any nonzero element c of $GF(p^r)$, and any nonnegative integers n and N,

$$W[P(x)(x^n - c)^N] \geq W[(x - c)^N]W[P(x) \bmod (x^n - c)].$$

Here $W[P(x)]$ denotes the Hamming weight of the polynomial $P(x)$, and we write $P(x) \bmod Q(x)$ for the remainder when $P(x)$ is devided by the polynomial $Q(x)$.

We shall describe a certain class of fixed nonsystematic convolutional codes of rate κ/v using the following notation, which generalizes the notation of Massey [3].

Manuscript received July 5, 1972; revised September 16, 1972.
The author was with the Laboratory for Communication Theory, Technical University of Denmark, Lyngby, Denmark. He is now with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Calif. 94720.

The information sequence with zeros inserted in the check positions has the D-transform

$$I(D) = i_0 + i_1 D + i_2 D^2 + \cdots + i_{\kappa-1} D^{\kappa-1} + i_v D^v + \cdots = \sum_{j=1}^{\kappa} D^{j-1} I_j(D^v).$$

In this paper a convolutional code is defined by a single generator polynomial $G(x)$ and the D-transform of the encoded sequence may be written as $T(D) = I(D)G(D)$.

We define the constraint length of the code as $n_A =$ degree $(G) + 1$. Thus the constraint length is the number of encoded symbols affected by a single information symbol.

A general fixed convolutional code of rate κ/v is defined by a generator matrix of the form [4]

$$\mathcal{G} = \begin{Vmatrix} G_{11}(D) & G_{12}(D) & \cdots & G_{1v}(D) \\ G_{21}(D) & G_{22}(D) & \cdots & G_{2v}(D) \\ \cdots & \cdots & \cdots & \cdots \\ G_{\kappa 1}(D) & G_{\kappa 2}(D) & \cdots & G_{\kappa v}(D) \end{Vmatrix}.$$

The subclass considered here has generator matrices

$$\mathcal{G}' = \begin{Vmatrix} G_1(D) & G_2(D) & \cdots & G_v(D) \\ DG_v(D) & G_1(D) & \cdots & G_{v-1}(D) \\ \cdots & \cdots & \cdots & \cdots \\ DG_{v-\kappa+2}(D) & DG_{v-\kappa+3}(D) & \cdots & G_{v-\kappa+1}(D) \end{Vmatrix}.$$

This class of convolutional codes is sufficiently large to allow a proof of the Gilbert lower bound on minimum distance by the following standard argument [5].

Consider a truncated encoded sequence T_m of length mv starting with a nonzero digit and a sequence of $m\kappa$ information symbols I_m also with a nonzero digit in the first position. (We recall that the minimum distance is defined as the minimum weight of any such T_m.) It is then easy to see that there is exactly one code in the class with constraint length $n_A \leq mv$ that produces an encoded sequence whose first mv digits are T_m as a response to an information sequence whose first $m\kappa$ digits are I_m. Consequently at least one code over $GF(q)$ with constraint length $\leq mv$ has minimum distance d_G satisfying

$$\sum_{i=0}^{d_G-1} \binom{mv}{i} (q-1)^i < q^{mv(1-\kappa/v)}. \tag{1}$$

A convolutional code is said to be catastrophic [6] if a nonpolynomial $I(D)$ can result in a polynomial $T(D)$ [4].

The free distance d_{free} of the convolutional code is the minimum of $W[T(D)]$ taken over all $I(D) \neq 0$. For non-

catastrophic codes, d_{free} may be taken as the minimum of $W[T(D)]$ over all polynomial $I(D) \neq 0$. The Gilbert bound (1) is clearly a lower bound on the free distance as well as the minimum distance.

We shall study convolutional codes generated by the generator polynomials $g(x)$ of cyclic codes. We shall write $h(x) = (x^n - 1)/g(x)$ for the generator of the cyclic dual code, d_g for the minimum distance of the original cyclic code, d_h for the minimum distance of the dual code, and n for the length of both cyclic codes.

The following result was derived in [1].

Theorem 2: If $g(x)$ generates a cyclic code over $GF(2^r)$ of odd length n, then for any positive integer m the rate $R = 1/v$ 2^r-ary convolutional code with $v = 2m$ defined by $G(D) = g(D)$ is noncatastrophic and has $d_{free} \geq \min\{d_g, 2d_h\}$.

In Section III we give a generalization of this theorem to fields of odd characteristic.

Since the duals of many good cyclic codes, notably BCH codes, have small minimum distances, it is desirable to obtain a bound on the free distance that depends only on d_g. The following definition will be useful in deriving such a bound.

Definition: Let p be a prime, n an integer relatively prime to p, and v an integer that divides n. The relation

$$\alpha \equiv \beta, \qquad \text{iff } \alpha^v = \beta^v$$

is an equivalence relation among the nth roots of unity in $GF(p^{rs})$. We shall say that the v equivalent nth roots of unity form a v-class or that they are v-equivalent.

If β is a primitive nth root of unity, $\gamma = \beta^{n/v}$ and α is an nth root of unity, the v-class that contains α is $\{\alpha, \alpha\gamma, \alpha\gamma^2, \cdots, \alpha\gamma^{v-1}\}$.

The definition of a convolutional code introduced in this section makes it possible to derive convolutional codes of rate κ/v, $\kappa \geq 1$, from cyclic codes in a natural way.

Theorem 3: If $g(x)$ generates a cyclic code over $GF(p^r)$ of length n relatively prime to p, v is any positive integer that divides n, and $g(x)$ has at most $v - \kappa$ v-equivalent roots, then the rate $R = \kappa/v$ convolutional code over $GF(p^r)$ generated by $G(x) = g(x)$ is noncatastrophic and has free distance $d_{free} \geq d_g$.

We base the proof on the following lemma.

Lemma 1: If κ or more v-equivalent nth roots of unity are roots of a polynomial of the form

$$P(x) = \sum_{j=1}^{\kappa} x^{j-1} P_j(x^v)$$

then all elements of the v-class are roots of $P(x)$.

Proof: Let α be a root of $P(x)$, that is

$$P_1(\alpha^v) + \alpha P_2(\alpha^v) + \cdots + \alpha^{\kappa-1} P_\kappa(\alpha^v) = 0. \qquad (2)$$

The component polynomials $P_j(x^v)$ have the same values for all v-equivalent values of x, and we may consequently interpret (2) as the equation

$$P_1 + zP_2 + z^2 P_3 + \cdots + z^{\kappa-1} P_\kappa = 0. \qquad (3)$$

If any of the P_j are nonzero, (3) has at most $\kappa - 1$ solutions for z. A polynomial that includes κ or more v-equivalent roots can thus divide $P(x)$ only if $P_j(\alpha^v) = 0$, for all j. But in this case all elements of the v-class are roots of $P(x)$.

Proof of Theorem 3: To show that the convolutional code generated by $g(x)$ is noncatastrophic, we must prove that a nonpolynomial information sequence

$$I(D) = \sum_{j=1}^{\kappa} D^{j-1} \frac{P_j(D^v)}{Q_j(D^v)}$$

cannot produce a polynomial output.

We notice that the least common multiple of the $Q_j(D^v)$ is a polynomial in D^v, $Q(D^v)$. Thus

$$I(D) = [1/Q(D^v)] \sum_{j=1}^{\kappa} D^{j-1} R_j(D^v) = R(D)/Q(D^v).$$

It may be assumed that any common factors of $R(D)$ and $Q(D^v)$, that are polynomials in D^v have been canceled out, but they may have other common factors. For the output $T(D) = G(D)I(D)$ to be a polynomial all factors of $Q(D^v)$ must be factors of either $G(D)$ or $R(D)$. But if β is a root of $Q(D^v)$, all elements of the v-class containing β are roots. However, we have assumed that at most $v - \kappa$ of these elements are roots of $G(D)$, and consequently at least κ elements must be roots of $R(D)$ for a nonpolynomial $I(D)$ to produce a polynomial $T(D)$. We can now apply Lemma 1 to show that this is possible only if all elements of the v-class are roots of $R(D)$, which contradicts the assumption that $R(D)$ and $Q(D^v)$ have no polynomial in D^v as a common factor.

To bound the free distance of the convolutional code, we first observe that since v divides n, $(D^n - 1)^t$ is a polynomial in D^v.

Consequently $I(D)/(D^n - 1)^t$ is a polynomial of the same form as $I(D)$ for any t such that $(D^n - 1)^t$ divides $I(D)$.

Since we have assumed that $g(x)$ has at most $v - \kappa$ v-equivalent roots and since all nth roots of unity are roots of $x^n - 1 = g(x)h(x)$, it follows that at least κ elements of each v-class are roots of $h(x)$. Hence Lemma 1 shows that $h(D)$ divides $I(D)$ only if $(D^n - 1)$ is a factor of $I(D)$. Repeating this argument we obtain $I(D) = (D^n - 1)^N p(D)$ where $h(D)$ does not divide $p(D)$. Consequently with the aid of Theorem 1 we have

$$W[T(D)] \geq W[(D-1)^N] W[p(D)g(D) \bmod (D^n - 1)] \geq d_g$$

This completes the proof of Theorem 3.

The condition on the roots of $g(x)$ may appear quite restrictive, but we shall find in Sections III and IV that it is satisfied in several interesting cases.

The binary codes derived from Theorem 3 do not include codes of rate $\frac{1}{2}$. The most interesting binary codes will be those of rates $\frac{1}{3}$ and $\frac{2}{3}$.

III. Convolutional Codes With $d_{free} = n_A$

Since $W[G(D)] \leq n_A$, we have for any convolutional code $d_{free} \leq n_A$. The codes for which $d_{free} = n_A$ may be viewed as being analogous to the maximum-distance

separable block codes [7, p. 309]. The Reed–Solomon codes [7, p. 310] are a class of cyclic maximum distance separable codes that exist over any finite field $GF(q)$.

Let $n = q - 1$ be the length of a Reed–Solomon code over $GF(q)$, and let ν divide n. The generator polynomial of the Reed–Solomon code of rate $k/n = \kappa/\nu$ may be taken as

$$g(x) = (x - 1)(x - \alpha)(x - \alpha^2)\cdots(x - \alpha^{n-k-1}),$$

where α is a primitive element of $GF(q)$. From the ν-class containing α^s, $0 \le s < n/\nu$, the elements

$$\alpha^s, \alpha^{s+n/\nu}, \alpha^{s+2n/\nu}, \cdots, \alpha^{s+(n/\nu)(\nu-\kappa-1)}$$

are roots of $g(x)$. Thus $g(x)$ has exactly $\nu - \kappa$ roots from each ν-class, and consequently the condition of Theorem 3 is satisfied.

Theorem 4: If $g(x)$ is the generator polynomial of a Reed–Solomon code of rate $k/n = \kappa/\nu$ over $GF(q)$, then the q-ary convolutional code of rate κ/ν generated by $G(x) = g(x)$ is noncatastrophic and has $d_{\text{free}} = n_A = n - k + 1$.

The rate $\frac{1}{2}$ codes obtained in [1] by applying Theorem 2 to the Reed–Solomon codes over $GF(2^r)$ of rates $\frac{1}{3}$ also have $d_{\text{free}} = n_A$, but they have constraint lengths about 50 percent greater than the codes of Theorem 4 for rates close to $\frac{1}{2}$. The proof of Theorem 2 in [1] relies heavily on the relation between ν and the characteristic of the field, and we could extend this theorem in an obvious way to codes of rate $1/p$ for fields of characteristic p. We prefer to prove a less obvious, but stronger, generalization of Theorem 2.

Theorem 5: If $g(x)$ generates a cyclic code of odd length n over $GF(q)$, $q = p^r$ and p an odd prime, then the rate $\frac{1}{2}$ convolutional code defined by $G(D) = g(D)$ is noncatastrophic and has $d_{\text{free}} \ge \min\{d_g, 2d_h\}$.

Proof: We shall write $\tilde{P}(x)$ for the polynomial whose roots are the negatives of the roots of $P(x)$. Thus $\tilde{P}(x) = \pm P(-x)$ and $W[\tilde{P}(x)] = W[P(x)]$. Any polynomial $Q(x^2)$ can be factored into a product of the form $P(x)\tilde{P}(x)$, since all roots occur in pairs $\pm\sqrt{r^2}$. In particular $x^{2n} - 1 = (x^n - 1)(x^n + 1)$, and if α is a root of $x^n - 1$, $-\alpha$ is a root of $x^n + 1$. For odd n, α and $-\alpha$ cannot both be roots of $x^n - 1$ and consequently both cannot be roots of $g(x)$.

Now $G(D) = G_1(D^2) + DG_2(D^2)$: thus if α is a common root of the code-generating polynomials $G_1(D^2)$ and $G_2(D^2)$, $-\alpha$ is also a common root. We conclude that gcd $\{G_1, G_2\} = 1$, and hence that the code is noncatastrophic [6].

For any polynomial $I(D) \neq 0$ we may write

$$\begin{aligned} T(D) &= I(D^2)g(D) = I_1(D)\tilde{I}_1(D)g(D) \\ &= P(D)g(D)^{i+1}\tilde{g}(D)^i h(D)^j \tilde{h}(D)^j, \end{aligned} \tag{4}$$

where $P(D)$ is not divisible by any of the polynomials $g(D)$, $\tilde{g}(D)$, $h(D)$, $\tilde{h}(D)$.

Suppose first that $i \ge j$. Then (4) becomes

$$T(D) = P(D)g(D)^{i-j+1}\tilde{g}(D)^{i-j}(D^{2n} - 1)^j$$

since $(\widetilde{D^n - 1}) = D^n + 1$. None of the roots of $\tilde{g}(x)$ are roots of $(x^n - 1)$, so $h(x)$ cannot divide $\tilde{g}(x)^{i-j}$. Applying Theorem 1 we find

$$\begin{aligned} &W[T(D)] \\ &\quad \ge W[(D-1)^j]W[P(D)g(D)^{i-j+1}\tilde{g}(D)^{i-j} \\ &\qquad\qquad \cdot \bmod (D^{2n} - 1)] \\ &\quad \ge W[P(D)g(D)^{i-j+1}\tilde{g}(D)^{i-j} \bmod (D^n - 1)] \ge d_g. \end{aligned} \tag{5}$$

Conversely, suppose $i < j$, then from (4) we have

$$T(D) = P(D)h(D)^{j-i-1}\tilde{h}(D)^{j-i}(D^n - 1)(D^{2n} - 1)^i.$$

Again applying Theorem 1 we get

$$\begin{aligned} W[T(D)] \ge W[(D-1)^i]W[P(D)h(D)^{j-i-1}\tilde{h}(D)^{j-1} \\ \cdot (D^n - 1) \bmod (D^{2n} - 1)]. \end{aligned}$$

Here we note that $\tilde{h}(x)(x^n - 1)$ is the generator of a cyclic code of length $2n$ and minimum distance $2d_h$, since $\tilde{h}(x)$ divides $x^n + 1$, and $W[\tilde{P}(x)] = W[P(x)]$. Consequently

$$W[T(D)] \ge 2d_h. \tag{6}$$

The theorem now follows from (5) and (6).

Theorem 5 is not applicable to generators of Reed–Solomon codes in their usual form, because these codes have even lengths when the characteristic of the field is odd. However, the maximum distance separable cyclic p-ary codes generated by $g(x) = (x - 1)^i$ [8], [9], [1] may be used in Theorem 5 to obtain good p-ary convolutional codes.

IV. Convolutional Codes over Small Fields

In this section, we shall study the restriction to smaller fields of the codes defined by Theorem 4. While the BCH block codes may be obtained simply as the subcodes of Reed–Solomon codes with coefficients in a subfield [10], we must here observe the extra condition imposed by Theorem 3. Thus from a (primitive or nonprimitive) BCH code of sufficiently high rate $R \ge \kappa/\nu$ we may construct a convolutional code of rate κ/ν with free distance lower-bounded by the minimum distance of the BCH code, but shorter constraint length $n_A \simeq n(1 - R)$.

The following theorem demonstrates that when convolutional codes are derived from certain nonprimitive BCH codes, the condition of Theorem 3 can be checked without a detailed inspection of the roots of the generator polynomial.

Theorem 6: Let $n = \mu\nu$ be a divisor of $q^{\mu-1} - 1$ and μ a prime that divides no number of the form $q^r - 1$ for $r < \mu - 1$. If $g(x)$ is the generator polynomial of a q-ary cyclic code of length n, at most $\nu - \kappa$ of the νth roots of unity in $GF(q^{\mu-1})$ are roots of $g(x)$ and at most $\nu - \kappa$

irreducible polynomials of degree $\mu - 1$ are factors of $g(x)$, then the convolutional code defined by $G(x) = g(x)$ is noncatastrophic and has free distance $d_{free} \geq d_g$.

Proof: It follows from the assumptions that the irreducible factors of $x^n - 1$ are [7, p. 101] v polynomials of degree $\mu - 1$, $(x - 1)$, and one or more minimal polynomials of the vth roots of unity, $\alpha^{\mu}, \alpha^{2\mu}, \cdots, \alpha^{(v-1)\mu}$, where α is a primitive nth root of unity in $GF(q^{\mu-1})$.

For two roots of the same irreducible polynomial to be v-equivalent, we must have

$$\alpha^{tq^i} = \alpha^{t+j\mu}, \qquad i < \mu - 1,$$

which implies

$$t(q^i - 1) \equiv j\mu \bmod n, \qquad i < \mu - 1$$

and since μ is a factor of n

$$t(q^i - 1) = j'\mu, \qquad i < \mu - 1. \tag{7}$$

Since μ does not divide $(q^i - 1)$, (7) is satisfied only if t is a multiple of μ. Thus we may conclude that while all roots of the minimal polynomials of the vth roots of unity belong to the v-class containing 1, the roots of each of the irreducible factors of degree $\mu - 1$ include exactly one element from each of the remaining v-classes. Consequently the condition of Theorem 3 is satisfied if at most $v - \kappa$ of the irreducible factors of degree $\mu - 1$ divide $g(x)$, and at most $v - \kappa$ vth roots of unity are roots of $g(x)$.

We apply Theorem 6 in the following two examples.

Example 1: Since $65 = 5 \cdot 13$ divides $2^{12} - 1$ and 13 divides no smaller number of the form $2^r - 1$, $x^{65} + 1$ has 5 irreducible factors of degree 12, and the other factors are $x + 1$ and the minimal polynomial of α^{13}. We obtain the convolutional codes listed in Table I. It is interesting to observe [11] that the code of rate $\frac{1}{5}$ may be improved by including the minimal polynomial of α^{13} rather than $x + 1$. This generator polynomial includes exactly 4 elements of each v-class.

Example 2: $20 = 4 \cdot 5$ divides $3^4 - 1$ and 5 divides no smaller number of the form $3^r - 1$. The irreducible factors of $x^{20} - 1$ are 4 polynomials of degree 4, $x + 1$, $x - 1$, and the minimal polynomial of α^5.

From the generator polynomial with roots $\alpha, \alpha^2, \alpha^4, \alpha^5, \alpha^{10}$ we obtain a ternary rate $\frac{1}{4}$ code with $n_A = 16$ and $d_{free} \geq 11$.

Unfortunately the applicability of Theorem 6 is rather limited, but it provides some insight into the relationship between the divisors of n and the maximal degree of the generator polynomials that satisfy the condition of Theorem 3. In general this relationship is more complex, but the best generator polynomials may be readily determined by successive calculation of the roots of the minimal polynomials.

In Table II a list of binary and ternary codes is given. The values of d_{free}/n_A for the long binary codes in this list are comparable to the asymptotic values of the Gilbert bound and the stronger lower bound on free distance for nonsystematic codes derived by Neumann [12]. The binary rate $\frac{2}{3}$ code with $n_A = 102$ has $d_{free}/n_A = 0.22$, whereas the value of Neumann's bound is only 0.12. The rate $\frac{1}{3}$ code with $n_A = 517$ has $d_{free}/n_A = 0.22$ where the values of Gilbert's and Neumann's bounds are 0.17 and 0.34.

TABLE I

κ/ν	Roots of $g(x)$	d_{free}	n_A
4/5	0,1	≥ 6	14
3/5	0,1,3	≥ 10	26
2/5	0,1,3,5	≥ 14	38
1/5	0,1,3,5,7	≥ 22	50
1/5	1,3,5,7,13	≥ 25	53

Convolutional codes derived from cyclic codes of length 65. The generator polynomial is the product of the minimal polynomials of α^i where α is a primitive 65th root of unity and i assumes the values listed under "roots of $g(x)$."

TABLE II

q	n	κ/ν	Roots of $g(x)$	d_{free}	n_A
2	129	2/3	0,1,3	≥ 10	30
2	129	1/3	0,1,3,5,7,9,11	≥ 26	86
2	255	2/3	-1,-3,0,1,3,5,7,9	≥ 16	58
2	255	1/3	-17,-15,...,-1,0,1,...,17	≥ 38	130
2	023	2/3	-9,-7,...,-1,0,1,...,7,9	≥ 22	102
2	023	1/3	-31,-29,...,0,...,77,79	≥ 116	517
3	242	1/2	-11,-10,...,0,...,10,11	≥ 26	82
3	244	3/4	0,1,2,4,5	≥ 14	42
3	244	2/4	0,1,2,4,...,13,14	≥ 32	102
3	244	1/4	0,1,2,...,17,19,20	≥ 44	144

Convolutional codes over $GF(q)$ derived from cyclic codes of length n. The generator polynomial is the product of the minimal polynomials of α^i where α is a primitive nth root of unity and i assumes the values listed under "roots of $g(x)$."

V. Asymptotically Good Periodic Convolutional Codes

In this section we shall construct for any R, $0 < R < 1$, a sequence of convolutional codes of rates $R_{n_A} \geq R$ for increasing n_A such that the free distances and constraint lengths satisfy

$$\lim_{n_A \to \infty} \inf d_{free}/n_A > 0. \tag{8}$$

In order to simplify the proof we shall consider only binary codes.

Let $g_1(x)$ be the generator polynomial of a convolutional code of rate $R_m = \kappa/\nu$ over $GF(2^m)$ derived from Theorem 4. We write the D-transform of the encoded sequence

$$T_1(D) = t_0 + t_1 D + t_2 D^2 + \cdots + t_u D^u + \cdots.$$

In a way similar to the construction in [2], define a time-varying convolutional code of rate $\kappa/2\nu$ by the encoded

sequence

$$T(D) = t_0 + t_0 D + t_1 D^2 + \alpha t_1 D^3 + \cdots + t_u D^{2u} + \alpha^u t_u D^{2u+1} + \cdots, \quad (9)$$

where α is a primitive element of $GF(2^m)$.

Alternatively we could have defined the convolutional code by two code-generating polynomials, the fixed polynomial $g_1(x)$ and a time-varying polynomial, but we prefer the simpler definition directly in terms of the encoded sequence (9).

We shall express the elements of $GF(2^m)$ as binary m-place vectors, and we may thus interpret the code defined by (9) as a time-varying binary convolutional code. The period of the code is $2m(2^m - 1)$ since $\alpha^{2^m-1} = 1$.

In our proof of theorem 3, we noted that the encoded sequence $T_1(D)$ could be written

$$T_1(D) = g_1(D)P(D)(D^n - 1)^N,$$

where $g_1(D)P(D)$ is not divisible by $(D^n - 1)$. We need more information about $T_1(D)$ than just the weight and, therefore, a few steps of the proof of Theorem 1 [1] follow.

In order to bound the Hamming weight of a polynomial of the form $q(x) \cdot (x^n + 1)^N$ we write

$$q(x) = q_0(x^n) + xq_1(x^n) + \cdots + x^{n-1}q_{n-1}(x^n).$$

Then

$$W[q(x)(x^n + 1)^N] = \sum_{i=0}^{n-1} W[q_i(x^n)(x^n + 1)^N]$$

and the theorem is proved by observing that for each term

$$W[q_i(x^n)(x^n + 1)^N] \geq W[q_i(1)]W[(x + 1)^N].$$

However, the weight of each of these terms is the weight of the symbols that are multiplied by α^i when we form the sequence (9), and consequently

$$W[P(D)g(D) \bmod (D^n - 1)] \geq d_g$$

is a lower bound on the number of nonzero symbols in $T_1(D)$ that are multiplied by distinct powers of α when $T(D)$ is formed.

We are now in a position to apply the lemma derived in [2].

Lemma 2: Let $o_2(L) \to 0$ as $L \to \infty$. Then for any γ, $0 < \gamma < 1$, and any δ, $0 < \delta < 1$, the total Hamming weight W of $M_L = [\gamma - o_2(L)](2^{L\delta} - 1)$ distinct nonzero binary L-tuples satisfies

$$W \geq \gamma L[H^{-1}(\delta) - o_1(L)](2^{L\delta} - 1).$$

Here H^{-1} denotes the inverse of the binary entropy function, and $o_1(L) \to 0$ as $L \to \infty$. We apply Lemma 2 with $L = 2m$, $\delta = \frac{1}{2}$, and $\gamma = 1 - R_m$. Hence

$$d_{\text{free}} \geq 2m(1 - R_m)[H^{-1}(\tfrac{1}{2}) - o_1(2m)](2^m - 1).$$

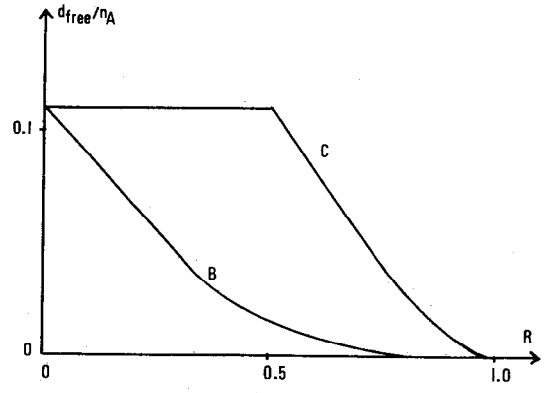

Fig. 1. A comparison of the bound on d_{free}/n_A for the convolutional codes constructed in Section V (C) and the bound obtained earlier for a class of block codes (B).

Now $n_A = 2m(2^m - 1)(1 - R_m)$, and consequently

$$\liminf_{n_A \to \infty} d_{\text{free}}/n_A \geq H^{-1}(\tfrac{1}{2}) \quad (10)$$

for any rate $R < \frac{1}{2}$.

As in [2] we obtain codes of higher rates by deleting the last s binary digits from each of the products $t_u\alpha^u$, and we shall write $[t_u\alpha^u]_s$ for the resulting $(m - s)$-place vector. We modify the encoded sequence to

$$T_s\colon t_0,[t_0]_s,t_1,[\alpha t_1]_s,\cdots,t_u,t_u[\alpha^u]_s,\cdots. \quad (11)$$

There are at least $(2^m - 1)(1 - R_m)2^{-s}$ distinct nonzero vectors of the form $\{t_u,[t_u\alpha^u]_s\}$, and we can again apply Lemma 2, now taking $L = 2m - s$, $\delta = (m - s)/(2m - s)$, and $\gamma = (1 - R_m)$. The total weight of the nonzero vectors is

$$W \geq 2^s(1 - R_m)(2m - s) \cdot \left[H^{-1}\left(\frac{m - s}{2m - s}\right) - o_1(m)\right](2^{m-s} - 1).$$

We note that the rate of this binary code is $R_{n_A} = R_m m/(2m - s)$, but the best bound is obtained for R_m close to 1, so that $R_{n_A} \simeq m/(2m - s)$. The constraint length is

$$n_A = (2m - s)(2^m - 1)(1 - R_m).$$

Thus

$$\liminf_{n_A \to \infty} d_{\text{free}}/n_A \geq H^{-1}(1 - R). \quad (12)$$

We combine (10) and (12) to obtain the following theorem.

Theorem 7: The binary periodic convolutional codes defined by (9) and (11) have free distances and constraint lengths satisfying

$$\liminf_{n_A \to \infty} d_{\text{free}}/n_A \geq \begin{cases} H^{-1}(\frac{1}{2}), & 0 < R < \frac{1}{2} \\ H^{-1}(1 - R), & \frac{1}{2} \leq R < 1. \end{cases}$$

The bound of Theorem 7 is plotted in Fig. 1 together with

the corresponding bound for block codes obtained in [2]. For rates greater than $\frac{1}{2}$ the bound obtained in Theorem 7 equals the Gilbert lower bound for block codes. By concatenating the convolutional codes of Theorem 4 with good fixed block codes of rates less than $\frac{1}{2}$ one could obtain a class of convolutional codes whose free distances are lowerbounded by the Gilbert bound for all rates $0 < R < 1$ [10], [13].

Acknowledgment

The author would like to thank Prof. J. L. Massey and the referees for several helpful suggestions.

References

[1] J. L. Massey, D. J. Costello, and J. Justesen, "Polynomial weights and code constructions," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 101–110, Jan. 1973.

[2] J. Justesen, "A class of constructive asymptotically good algebraic codes," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 652–656, Sept. 1972.

[3] J. L. Massey, "Majority decoding of convolutional codes," Res. Lab. Electron. Quart. Prog. Rep. 64, Massachusetts Inst. Technol., Cambridge, pp. 183–188, Jan. 15, 1962.

[4] G. D. Forney, Jr., "Convolutional codes I: Algebraic structure," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 720–738, Nov. 1970.

[5] J. L. Massey, "Some algebraic and distance properties of convolutional codes," in H. B. Mann, *Error Correcting Codes*. New York: Wiley, 1968.

[6] J. L. Massey and M. K. Sain, "Inverses of linear sequential circuits," *IEEE Trans. Comput.*, vol. C-17, pp. 330–337, Apr. 1968.

[7] E. R. Berlekamp, *Algebraic Coding Theory*. New York: McGraw-Hill, 1968.

[8] E. F. Assmus and H. F. Mattson, Jr., "New 5-designs," *J. Comb. Theory*, vol. 6, pp. 122–151, 1969.

[9] S. D. Berman, "On the theory of group codes," *Kibernetika*, vol. 3, No. 1, pp. 31–39, 1967.

[10] G. D. Forney, Jr., *Concatenated Codes*. Cambridge, Mass.: M.I.T. Press, 1966.

[11] C. L. Chen, "Computer results on the minimum distance of some binary cyclic codes," *IEEE Trans. Inform. Theory* (Corresp.), vol. IT-16, pp. 359–360, May 1970.

[12] Neumann, B., "Distance properties of convolutional codes," M.S. thesis, Massachusetts Inst. Technol., Cambridge, Aug. 1968.

[13] V. V. Zyablov, "On estimation of complexity of construction of binary linear concatenated codes," *Probl. Peredach. Inform.*, vol. 7, pp. 5–13, 1971.

Correspondence

Time, Frequency, Sequency, and their Uncertainty Relations

JUDEA PEARL

***Abstract*—We study the form assumed by the classical time-frequency uncertainty relations in discrete as well as nontrigonometric spectral analysis. In particular we find that if an N-sample time signal is to contain a fraction γ of its energy in T consecutive samples, then the minimum number of frequency components containing that same energy fraction must be greater than $N/T(2\gamma - 1)^2$. It is also found that the discrete Walsh transform permits greater energy concentration (less uncertainty) than the discrete Fourier transform.**

I. Introduction

It is well known that one cannot simultaneously confine a function $f(t)$ and its Fourier transform $F(\omega)$ without limit. This phenomenon is common to both continuous- and discrete-time functions. The most familiar form of the time-frequency uncertainty is the one leading to the Heisenberg *uncertainty principle*, stating that if we measure the time spread T of $f(t)$ by

$$T^2 = \frac{\int_{-\infty}^{\infty} (t - t_o)^2 |f(t)|^2 \, dt}{\int_{-\infty}^{\infty} |f(t)|^2 \, dt} \tag{1}$$

and the frequency spread Ω of $F(\omega)$ by

$$\Omega^2 = \frac{\int_{-\infty}^{\infty} (\omega - \omega_o)^2 |F(\omega)|^2 \, d\omega}{\int_{-\infty}^{\infty} |F(\omega)|^2 \, d\omega} \tag{2}$$

Manuscript received August 30, 1971; revised June 9, 1972.
The author is with the School of Engineering and Applied Science, University of California, Los Angeles, Calif. 90024.

then, for any choice of t_o and ω_o,

$$\Omega T \geq \tfrac{1}{2} \tag{3}$$

with equality holding if $f(t)$ is a Gaussian wave packet.

This study was motivated by the following dilemma: while we recognize that a similar limitation also exists in discrete time (one cannot simultaneously confine a sampled-time function and its discrete Fourier transform without limit), (3) does not seem to capture this limitation; in an N-dimensional space T can be made 0, Ω must remain finite, and so the product ΩT can be made as small as one desires. We therefore seek a more refined version of the uncertainty principle that lends itself to extension over finite-dimensional vector spaces.

An early work by Fuchs [2] (1956) provides the sought-for extension (unfortunately the proof is unpublished). Fuchs considers two arbitrary subsets S_T and S_Ω of finite measure in the time and frequency spaces, respectively. He proves that if the energy fraction in the time subset S_T is α^2, and the energy fraction in the frequency subset S_Ω is β^2, then

$$\begin{aligned} \beta^2 &\leq 1, && \alpha^2 \leq \lambda_o \\ &\leq \alpha\lambda_o^{1/2} + [(1 - \alpha^2)(1 - \lambda_o)]^{1/2}, && \alpha^2 > \lambda_o, \end{aligned} \tag{5}$$

where λ_o is the largest eigenvalue of the equation

$$(2\pi)^{-N} \int_{\omega \in S_\Omega} y(\omega) \int_{u \in S_\Omega} \exp[-i(t - \omega)u] \, du \, d\omega = \lambda y(t). \tag{6}$$

Equations (4) and (5) give the maximum frequency concentration β^2 that can be achieved with given time concentration α^2, time band S_T, and frequency band S_Ω.